

Objective reduction based on nonlinear correlation information entropy

Wang, Handing; Yao, Xin

DOI:

[10.1007/s00500-015-1648-y](https://doi.org/10.1007/s00500-015-1648-y)

License:

Creative Commons: Attribution (CC BY)

Document Version

Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Wang, H & Yao, X 2016, 'Objective reduction based on nonlinear correlation information entropy', *Soft Computing*, vol. 20, no. 6, pp. 2393–2407. <https://doi.org/10.1007/s00500-015-1648-y>

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Objective reduction based on nonlinear correlation information entropy

Handing Wang¹ · Xin Yao²

Published online: 17 March 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract It is hard to obtain the entire solution set of a many-objective optimization problem (MaOP) by multi-objective evolutionary algorithms (MOEAs) because of the difficulties brought by the large number of objectives. However, the redundancy of objectives exists in some problems with correlated objectives (linearly or nonlinearly). Objective reduction can be used to decrease the difficulties of some MaOPs. In this paper, we propose a novel objective reduction approach based on nonlinear correlation information entropy (NCIE). It uses the NCIE matrix to measure the linear and nonlinear correlation between objectives and a simple method to select the most conflicting objectives during the execution of MOEAs. We embed our approach into both Pareto-based and indicator-based MOEAs to analyze the impact of our reduction method on the performance of these algorithms. The results show that our approach significantly improves the performance of Pareto-based MOEAs on both reducible and irreducible MaOPs, but does not much help the performance of indicator-based MOEAs.

Keywords Multi-objective optimization · Objective reduction · Nonlinear correlation information entropy ·

Communicated by V. Loia.

✉ Handing Wang
wanghanding@163.com

Xin Yao
x.yao@cs.bham.ac.uk

¹ Key Lab of Intelligent Perception and Image Understanding of Ministry of Education, International Research Center of Intelligent Perception and Computation, Xidian University, Xi'an 710071, China

² CERCIA, School of Computer Science, The University of Birmingham, Birmingham B15 2TT, UK

Multi-objective evolutionary algorithm · Dimension reduction

1 Introduction

Loosely speaking, a many-objective optimization problem (MaOP) (Khare et al. 2003; Praditwong and Yao 2007) is a special kind of multi-objective optimization problem (MOP) with more than three objectives (Fleming et al. 2005; Hughes 2007). The large number of objectives in many-objective optimization brings many challenges to multi-objective evolutionary algorithms (MOEAs). Most solutions in the population of an MaOP are non-dominated (Ishibuchi et al. 2008), thus, the selection mechanism based on the Pareto dominance is less effective. Pareto-based MOEAs such as NSGA-II (Deb et al. 2002a) fail to solve MaOPs (Purshouse and Fleming 2003; Hughes 2005; Wagner et al. 2007; Khare et al. 2003). Without an effective dominance relation, MOEAs are unable to provide promising search directions (Deb and Jain 2014). Moreover, the growing number of objectives increases the computational complexity of Pareto-based MOEAs. Although the non-dominated rank sort (Deb and Tiwari 2005), deductive sort (McClymont and Keedwell 2012), and corner sort (Wang and Yao 2014) have been proposed to reduce that complexity, the progress is still unsatisfactory.

To overcome the difficulty of MaOPs, the existing research can be divided into five classes:

- *Dominance relation modification* The Pareto dominance is ineffective for MaOPs. Much work aims to modify the original dominance relation (Köppen et al. 2005; Kukkonen and Lampinen 2007; Sato et al. 2007; Farina and

Amato 2002; Dai et al. 2015), but their performance is still less than satisfactory.

- *Decomposition-based MOEAs* The main idea is to solve MaOPs by aggregation functions with a series of weight vectors to obtain several single-objective optimization problems (Zhang and Li 2007; Ma et al. 2014), but it suffers from poor performance on MaOPs with highly correlated objectives (Ishibuchi et al. 2009) because of the unsuitable arrangement of weight vectors (Ishibuchi et al. 2011a,b).
- *Indicator-based MOEAs* With a metric as a single objective, indicator-based MOEAs can avoid employing the Pareto dominance (Zitzler and Künzli 2004; Bader and Zitzler 2011; Wagner et al. 2007; Gong et al. 2014). However, $I_{\varepsilon+}$ in IBEA Zitzler and Künzli (2004) and I_H in HypE Bader and Zitzler (2011) provide unsatisfactory diversity on the Pareto front (PF) (Hadka and Reed 2012).
- *Incorporation with decision makers* Usually, decision makers do not need all the optimal solutions of MaOPs (Cvetkovic and Parmee 2002). They can input their interested regions or preferences to obtain parts of the non-dominated solution set (Sindhya et al. 2011; Ben Said et al. 2010; Koksalan and Karahan 2010; Wang et al. 2013a; Kim et al. 2012; Karahan and Koksalan 2010; Giagkiozis and Fleming 2014). Additionally, decision makers have different targets for different objectives and multi-target search was employed (Wang et al. 2013b).
- *Objective reduction* For some MOPs, unnecessary objectives can be ignored without changing their Pareto sets (Gal and Hanne 1999). Thus, the difficulty caused by a large number of objectives of a MaOP can be reduced (Fonseca and Fleming 1995; Coello Coello 2005; Deb 2001), and the existing Pareto-based MOEAs for MOPs with low-dimensional objectives can be used.

Objective reduction aims to make the problems with redundant objectives easier to solve by the existing MOEAs. The basic goal of objective reduction is to select the smallest set of objectives without changing the Pareto set of the original problem (Gal and Hanne 1999). In Brockhoff and Zitzler (2009), there is a related, but different explanation of this aim through the correlation among objectives, i.e., trying to obtain the smallest set of conflicting objectives. Based on different understanding of objective reduction, the existing objective reduction techniques can be divided into three classes:

- *Dominance relation preservation-based objective reduction* It is based on a measure for the changes of the dominance structure, which obtains a minimum subset of objectives with the preserved dominance relation (Brockhoff and Zitzler 2006). An additional term δ is adopted to measure the difference between the dominance struc-

tures of two subsets. However, the technique can only be applied to the linear objective reduction.

- *Pareto corner search* The Pareto corner search evolutionary algorithm (PCSEA) (Singh et al. 2011) is a newly proposed objective reduction approach. It only searches the corners of PFs. Then, it uses the obtained solutions to analyze the relation among objectives. Finally, it outputs a subset of non-correlated objectives. PCSEA is an off-line method.
- *Machine learning-based objective reduction* As the process of objective reduction can be seen as feature selection, this method focuses on the objectives with negative correlation and uses an improved correlation matrix of objectives to measure the conflict degree of two objectives (López Jaimes et al. 2008). With the obtained correlation matrix as distances, the method divides those objectives into neighborhoods. Then, it adopts a q -neighbor structure to select objectives. However, q has to be set in advance. Other machine learning techniques for dimension reduction, such as principal component analysis (PCA) and maximum variance unfolding (MVU), have also been applied to objective reduction (Saxena and Deb 2007; Deb and Saxena 2005). These objective reduction methods use machine learning techniques to select conflicting objectives according to the correlation information (the correlation matrix and correntropy matrix, for instance).

The aforementioned objective reduction approaches have their disadvantages. For example, both approaches in Brockhoff and Zitzler (2006) and PCSEA are off-line approaches for supporting decision makers after running MOEAs. This paper focuses on online objective reduction approaches. Although interpreting objective reduction through the correlation (Brockhoff and Zitzler 2009) is not exactly the same as the original definition (Gal and Hanne 1999), it covers the majority of cases in practice and is easy to apply to online objective reduction approaches. Therefore, we follow this interpretation of objective reduction and use the non-dominated population in every generation as the learning dataset to identify the redundant objectives in this paper.

Objectives that can be reduced are either linearly or nonlinearly correlated, mostly nonlinearly correlated (Saxena et al. 2013). However, the majority of the existing approaches use linear statistical tools to measure both linear and nonlinear correlation. In such cases, nonlinear correlation would be weakened by the linear description, which misleads the reduction.

In this paper, we use the same measurement for both linear and nonlinear correlation; thus, the performance of online objective reduction approaches can be improved. We find NCIE (Wang et al. 2005) to be a very robust measure for both linearly and nonlinearly correlated datasets, which has

been applied to the analysis of neurophysiological signals (Pereda et al. 2005), the quantification of the dependence among noisy data (Khan et al. 2007), etc. Therefore, we adopt NCIE as a correlation measure in objective reduction and study its impact on online objective reduction approaches.

The rest of the paper is organized as follows. We first show different cases of redundant objectives in MOPs in Sect. 2. In Sect. 3, NCIE is introduced. In Sect. 4, our approach will be described in detail. Section 5 reports the experimental results, in which the behavior of our approach is analyzed and discussed. Finally, Sect. 6 gives the conclusion and points out the future work.

2 Conflicting and redundant objectives

2.1 Conflicting objectives

Simply, the conflict between two objectives means that the improvement on one objective would deteriorate the other objective. The conflict might be global or local (the range of conflict) (Freitas et al. 2013), and linear or nonlinear (the structure of correlation) (Saxena et al. 2013).

2.2 Redundant objectives

If there is no conflict between two objectives, one of them can be viewed as a redundant objective for this MOP. Generally, the redundant objectives in an MOP are defined as the objectives that can be ignored without changing the structure of its original PF (Gal and Hanne 1999).

2.3 Reducible MOPs

Many-objective optimization problems (MOPs) with redundant objectives are reducible MOPs, which can be applied objective reduction techniques. If such MOPs can be reduced to MOPs with low-dimensional objectives, existing MOEAs can be used.

The above definition is not strictly mathematical. However, as Brockhoff and Zitzler (2006) mentioned, the existing literature has not clarified two main problems for objective reduction. One is the effect of objective reduction on dominance, and the other is the evaluation of the subset of objectives after reduction.

To model objective reduction mathematically, the redundant objectives are considered as the objectives positively correlated to some other objectives in the MOP (Brockhoff and Zitzler 2009). Actually, this transformation is not strictly equivalent. Table 1 shows some MOPs with a redundant objective f_3 . Their parallel coordinate graphs are shown in Fig. 1. In Cases 1 and 2, f_3 is positively correlated to f_1 , which can be reduced. In Case 3, f_3 is constant and non-

Table 1 Examples of minimal MOPs with redundant objectives, where $x \in [0, 1]$

Case 1	$f_1 = x_1$ $f_2 = 1 - x_1$ $f_3 = 2f_1$ $PF: f_1 + f_2 = 1, f_3 = 2f_1$
Case 2	$f_1 = x_1$ $f_2 = 1 - x_1$ $f_3 = \sin(0.5\pi f_1)$ $PF: f_1 + f_2 = 1, f_3 = \sin(0.5\pi f_1)$
Case 3	$f_1 = x_1$ $f_2 = 1 - x_1$ $f_3 = 1$ $PF: f_1 + f_2 = 1, f_3 = 1$
Case 4	$f_1 = x_1x_2(1 + x_3^2)$ $f_2 = x_1(1 - x_2)(1 + x_3^2)$ $f_3 = \begin{cases} (1 - x_1)(1 - x_2)(1 + x_3^2), & x_3 \neq 0 \\ 0, & x_3 = 0 \end{cases}$ $PF: f_1 + f_2 = 1, f_3 = 0$

correlated to any objective. In Case 4, f_3 is in conflict with f_1 and f_2 in most parts (locally), but it does not contribute to the PF structure, because f_3 is constant and non-correlated to any objective on the PF. However, Cases 3 and 4 cannot be covered by the above definition. Case 3 is special in the real world, and Case 4 is hard to be detected during the search. Therefore, we only focus on Cases 1 and 2 in this paper.

In Cases 1 and 2, the linear and nonlinear correlation are both important in objective reduction. However, the majority of the existing approaches employ linear tools to describe all the scenarios (Deb and Saxena 2005), which results in poor performance for nonlinear correlation. Comparing Cases 1 and 2, f_3 is a redundant objective for f_1 . In Case 1, f_3 is linearly correlated to f_1 , but nonlinearly correlated to f_1 in Case 2. If we use linear tools to evaluate the correlation degree, the obtained conflict degree in Case 2 is smaller than that in Case 1. It is obviously less reasonable. That is the reason why we use NCIE to capture a more general correlation for objective reduction.

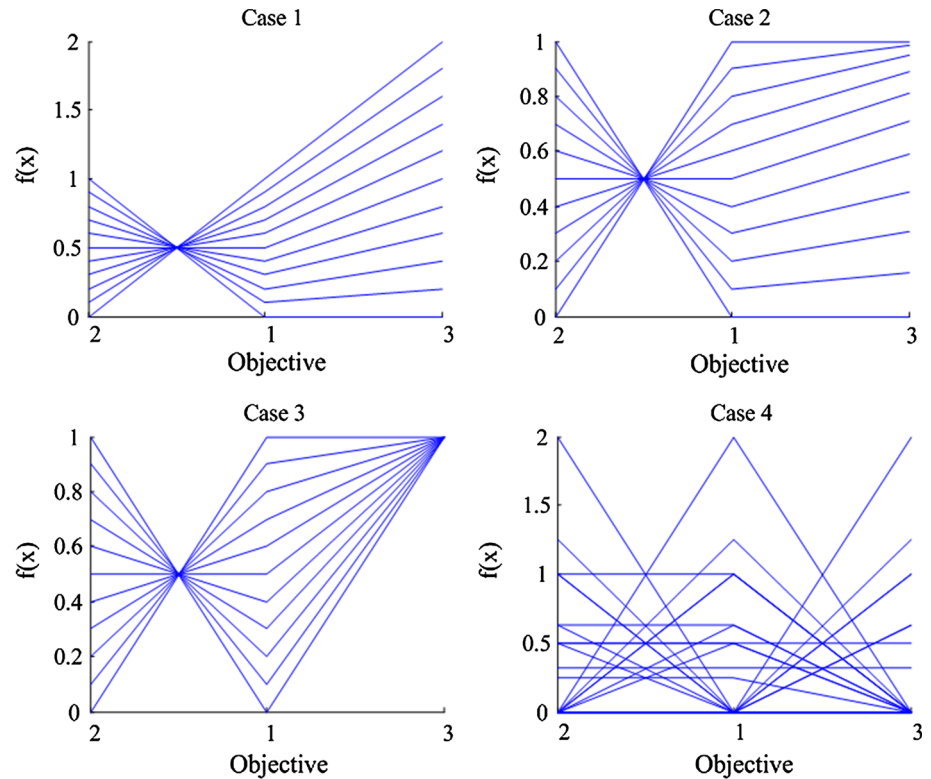
3 Nonlinear correlation information entropy

Mutual information entropy is a kind of generalized correlation; it is sensitive to different kinds of relation, which is shown in Eq. (1) (Maes et al. 1997),

$$I(X, Y) = H(X) + H(Y) - H(X, Y), \quad (1)$$

where X (with domain of L possible values) and Y (with domain of M possible values) are two discrete random variables, $H(X)$ is the information entropy of X , which is defined as Eq. (2). $H(X, Y)$ is the joint entropy of X and Y shown as Eq. (3). In Eq. (2), p_i is the probability of X with the i th

Fig. 1 Parallel coordinate graphs of Cases 1–4



value. Similarly, p_{ij} is the probability of X with the i th value and Y with the j th value in Eq. (3).

$$H(X) = - \sum_{i=1}^L p_i \ln(p_i). \quad (2)$$

$$H(X, Y) = - \sum_{i=1}^L \sum_{j=1}^M p_{ij} \ln(p_{ij}). \quad (3)$$

The authors in Wang et al. (2005) proposed a new non-linear correlation information entropy for multi-variable analysis. The results show that the new entropy quantizes the correlation in $[0, 1]$ for both linear and nonlinear cases (Wang et al. 2005).

Nonlinear correlation information entropy (NCIE) firstly divides variables X and Y into b rank grids. Then, the probabilities can be sampled by the counts in those grids. Thus, p_{ij} in the ij th grid can be calculated. The joint entropy is shown in Eq. (4),

$$H^r(X, Y) = - \sum_{i=1}^b \sum_{j=1}^b \frac{n_{ij}}{N} \log_b \left(\frac{n_{ij}}{N} \right), \quad (4)$$

where N is the size of the dataset, n_{ij} is the number of samples distributed in the ij th rank grid, and b is set to $\lceil \sqrt{N} \rceil$. NCIE is shown in Eq. (5), where $H^r(X)$ is the revised entropy of X as Eq. (6). Thus, the only parameter b is set self-adaptively,

which makes NCIE parameter self-adaptive. NCIE can also be calculated by a simple formula as Eq. (7).

$$\text{NCIE}(X, Y) = H^r(X) + H^r(Y) - H^r(X, Y). \quad (5)$$

$$H^r(X) = - \sum_{i=1}^b \frac{n_i}{N} \log_b \left(\frac{n_i}{N} \right). \quad (6)$$

$$\text{NCIE}(X, Y) = 2 + \sum_{i=1}^{b^2} \frac{n_i}{N} \log_b \left(\frac{n_i}{N} \right). \quad (7)$$

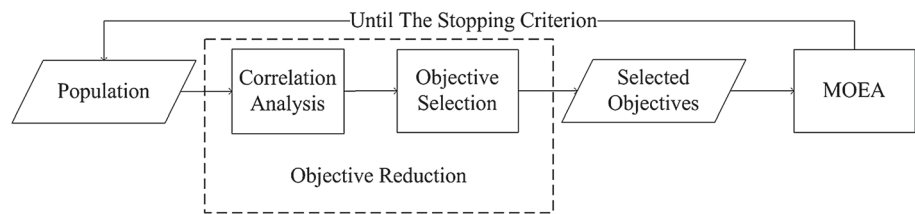
Based on the NCIE matrix $R^N = \{\text{NCIE}_{ij}\}$, ($1 \leq i \leq K$, $1 \leq j \leq K$), the relation among K variables can be analyzed.

4 Objective reduction based on nonlinear correlation information entropy

4.1 Basic idea

Our proposed method uses NCIE as a metric to reduce redundant objectives, whose flowchart is shown in Fig. 2. The proposed method first analyzes the correlation of objectives using the non-dominated population as its dataset. Based on the correlation of objectives, the method obtains a subset of conflicting objectives for MOEAs. Then, MOEAs only focuses on this subset of objectives, which is updated by the objective reduction approach in every generation.

Fig. 2 Flowchart of the proposed objective reduction approach embedded in an MOEA



The correlation analysis and objective selection are two key steps in an objective reduction approach. For correlation analysis, a majority of the existing approaches are based on the correlation matrix, which is only used for the linear correlation measure. As NCIE can handle both the linear and nonlinear correlation, we adopt it to measure the correlation in our approach. For objective selection, we abandon those common techniques in the existing approaches (such as PCA and feature selection) and design a straightforward method to select conflicting objectives (explained in Sect. 4.3).

The NCIE-based correlation analysis is based on the non-dominated population in every generation; thus, the conflict between objectives are local rather than global (López Jaimes et al. 2014). As Sect. 3 shows, the conflicts might be local in some cases. Thus, our proposed method could reduce some non-globally redundant objectives but locally-redundant objectives. During the execution of MOEAs, the conflict degree would be updated by the value of NCIE. In short, the basic idea of our approach is to keep the most conflicting objectives and omit the most positively correlated objectives in the NCIE matrix during run time.

4.2 Correlation analysis

Although NCIE can describe both the linear and nonlinear correlation between objectives, it cannot describe their conflicting relation. NCIE cannot be used directly in its original version for our aim. In view of this, NCIE is modified by adding the information of covariance. Covariance is valued in $[-1, 1]$, whose sign describes whether two variables are in conflict. The modified NCIE is shown in Eq. (8), where cov_{ij} is the ij th element in the correlation matrix.

$$R^N = \{Sgn(cov_{ij})NCC_{ij}\}, \quad (1 \leq i, j \leq K). \quad (8)$$

In the modified NCIE, the sign is from covariance, whose role is to show whether two objectives are in conflict. The modified NCIE can describe the conflict degree. If the modified NCIE of two objectives is a large positive value, the two objectives are highly positively correlated. If the modified NCIE of two objectives is a large negative value, the two objectives are highly conflicted. If two objectives have a modified NCIE around zero, they are not correlated. In this case, the sign of the modified NCIE is not very important, because the difference between the values with different

signs is small. With the modified matrix, we can use either a threshold or a classification method to determine the correlation degree of two objectives.

4.3 Objective selection

With the modified NCIE matrix, our approach selects the most conflicting objectives for MOEAs. Our approach is applied in every generation of MOEAs to update the correlation information among objectives. The details are shown in Algorithm 1, where S_r is the selected objective set and S_t is a temporary set. After the calculation of the modified NCIE matrix, our approach selects the most conflicting objective, which is the objective with the largest absolute sum of its negative NCIEs to other objectives. Then, it omits the objectives that are positively correlated to the selected objective. Finally, our approach outputs the selected objectives. In the process of omitting objectives, a threshold T is applied to determine whether two objectives are positively correlated. The effect of T is analyzed in Sect. 5.2.3.

Algorithm 1 Pseudo code of the objective selection in our approach.

```

1: Input:  $R^N$ -modified NCIE matrix of the current population (calculated as Sect. 4.2),  $m$ -number of objectives.
2: Output:  $S_r$ -selected objective set.
3: Parameter:  $T$ -threshold.
4:  $S_t = [1 : m]$ ,  $S_r = \emptyset$ .
5: While  $S_t \neq \emptyset$ 
6:   If all the elements in  $R^N$  are positive,
7:      $J = \operatorname{argmax}(\sum(R^N(1 : m, j)))$ . Find the most representative objective
8:   Else
9:      $J = \operatorname{argmin}(\sum(R^N(i, j)))$ , where  $R^N(i, j) < 0$  and  $1 \leq i \leq m$ . Find the most conflicting objective with remaining objectives
10:  End
11: Move  $f_J$  from  $S_t$  to  $S_r$ .
12:  $F = \{f_j | R^N(J, j) > T \max(R^N(J, j)) (f_j \in S_t)\}$ . Find set  $F$  with the objectives correlated to  $f_J$ .
13: Delete set  $F$  from  $S_t$ .
14: End
  
```

To show the process of our objective selection method, we take a modified NCIE matrix on DLTZ5(2,5) (Deb et al. 2002b) in Table 2 as an example (T is set as 0.3). Firstly, S_t is $\{f_1, f_2, f_3, f_4, f_5\}$, and S_r is \emptyset . Objective f_5 is selected

Table 2 Example of the modified NCIE matrix on DLTZ5(2,5)

	f_1	f_2	f_3	f_4	f_5
f_1	1.0000	0.4959	0.4244	0.5348	-0.3552
f_2	0.4959	1.0000	0.3972	0.4686	-0.3381
f_3	0.4244	0.3972	1.0000	0.4765	-0.4352
f_4	0.5348	0.4686	0.4765	1.0000	-0.4488
f_5	-0.3552	-0.3381	-0.4352	-0.4488	1.0000
$\sum \text{NCIE} < 0$	-0.3552	-0.3381	-0.4352	-0.4488	-1.5773

($S_r = \{f_5\}$, $S_t = \{f_1, f_2, f_3, f_4\}$), because it has the largest absolute sum of its negative NCIEs to other objectives (f_5 has the most conflicting degree with other objectives). There is no objective positively correlated to f_5 ; thus, there is not a redundant objective with f_5 in the remaining objectives. Then, objective f_4 is selected ($S_r = \{f_5, f_4\}$, $S_t = \{f_1, f_2, f_3\}$), because it has the largest absolute sum of NCIEs to other objectives. Objectives f_1, f_2, f_3 are omitted ($S_t = \emptyset$); they are all positively correlated to f_4 (not in conflict with f_4) as redundant objectives, because $R_{(4,1)}^N > T$, $R_{(4,2)}^N > T$, and $R_{(4,3)}^N > T$. Finally, our approach obtains the reduced objective set $\{f_5, f_4\}$, which represents the main conflict in DLTZ5(2,5).

Our approach is different from the approach that outputs a fixed number of objectives (Deb and Saxena 2005). It selects different numbers of objectives according to the situation of the current population, which is more robust for different problems.

4.4 Classification for correlated and non-correlated objectives

Parameter T ($[0, 1]$) is the threshold to determine the correlation degree between objectives. It is important to use a suitable T , because if T is too large, some redundant objectives may be regarded as non-reducible, or some conflicting objectives would not be retained. It is difficult to set T manually in advance without any knowledge of the optimization problem. Actually, the whole issue should be regarded as a classification problem that separates the objectives correlated to objective f_i from those non-correlated to objective f_i . We avoid the manual setting of T by Algorithm 2. As the clustering problem is a one-dimensional problem of a small size, any clustering technique can fulfill the task of cutting across the least dense area between the two clusters. In this paper, we use K-means for clustering. In practice, there is the situation that all the objectives are non-correlated or correlated to f_i . Therefore, we add two virtual values 0 and 1 during the clustering to handle such cases.

Algorithm 2 Pseudo code of classifying objectives.

```

1: Input:  $R$ -modified NCIE of the remaining objectives to objective  $f_i$ 
   ( $R = R^N(i, j)$ , ( $f_j \in S_t$ ) and ( $f_i \in S_r$ )),  $m$ -number of objectives.
2: Output:  $R'$  with classification for correlated and non-correlated
   objectives.
3: Add values 0 and 1 (the boundary of non-correlated and correlated)
   to  $R^N$  as  $R^*$ .
4: Sort  $R^*$  to  $R'$  in an ascending order.
5: Classify  $R'$  into two clusters ( $R'[1 : k]$  and  $R'[k + 1 : \text{end}]$ ). In other
   words, cut across the least dense area between the two clusters.
6:  $R'[1 : k]$  without 0 is the cluster of objectives non-correlated to  $f_i$ ,
    $R'[k + 1 : \text{end}]$  without 1 is the cluster of objectives correlated to
    $f_i$ .

```

4.5 Computational complexity

For an m -objective problem with a solution set of size N (in most cases, N is larger than m), the NCIE matrix calculation in the correlation analysis has $O(m^2N)$ complexity, and the objective selection has $O(m^2)$ complexity. Therefore, the total complexity of our method is $(O(m^2N))$ per generation.

5 Experimental studies

5.1 Test problems, metrics, and settings

As the DTLZ problems (Deb et al. 2002b) and WFG3 (Huband et al. 2006) are MOPs with different numbers of objectives, we adopt them as the test problems in our experiments. Among these test problems, DTLZ1-4 are irreducible, and DTLZ5 (Deb and Saxena 2005) and WFG3 are reducible. DTLZ5(I, M) is an M -objective problem with I conflicting objectives.

We use IGD (the average distance from the true PF to the obtained PF) (Van Veldhuizen and Lamont 1998; Zhang et al. 2008) to evaluate both convergence and diversity in our experiments. The calculation of IGD is shown in Eq. (9), where PF_{true} is a reference set that is uniformly sampled from the true PF. Most of the test problems in this paper can be reduced, whose PFs are degraded on their dimensions, and we uniformly sample PF_{true} in its objective-reduced space to guarantee the accuracy of IGD. Taking DTLZ5(2, M) as an example, it can be reduced to $\{f_{M-1}, f_M\}$; we sample PF_{true} in the space of f_{M-1} and f_M and then the values of other objectives can be calculated. It is worth noting that PF_{true} and $\text{PF}_{\text{obtained}}$ in IGD are both of their full dimensions rather than that after objective reduction to guarantee fair comparison:

$$\text{IGD}(\text{PF}_{\text{true}}, \text{PF}_{\text{obtained}}) = \frac{\sum_{v \in \text{PF}_{\text{true}}} d(v, \text{PF}_{\text{obtained}})}{|\text{PF}_{\text{true}}|}. \quad (9)$$

All the algorithms in the following subsections are repeated for 30 independent runs and stop after 200 generations (SBX

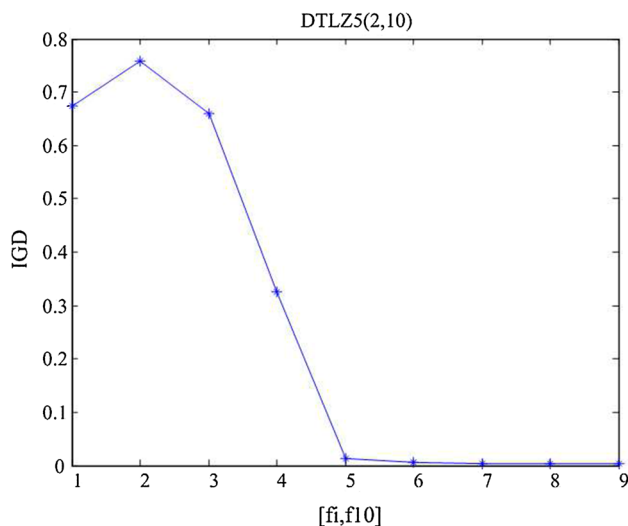


Fig. 3 Average IGD values of NSGA-II on DTLZ5(2,10) with only two selected objectives $\{f_i, f_{10}\} (1 \leq i \leq 9)$

crossover ($\eta = 15$) with probability 1 and polynomial mutation ($\eta = 15$) with probability 0.1). The population size is 200 (due to larger numbers of objectives in the test problems of following subsections, we will enlarge the size of the population later). The machine characteristics in our experiments are a T5470 1.6-GHz CPU and 1G RAM.

5.2 Experiments on analysis of components

5.2.1 Characteristics of DTLZ5

As we know, the objectives of DTLZ5(2, M) can be reduced to a two-objective problem with only two selected objectives $\{f_i, f_M\} (1 \leq i \leq M - 1)$. Further, $\{f_{M-1}, f_M\}$ is the best reduction result. To demonstrate this, we compare NSGA-II on DTLZ5(2,10)s with only two selected objectives $\{f_i, f_{10}\}$. The average IGD values of 30 independent runs are shown in Fig. 3. Although DTLZ5(2,10)s with $\{f_i, f_{10}\}$ all have two objectives, the difficulties are not at the same level. In Fig. 3, we find the increasing i decreases the difficulty of DTLZ5(2,10), which can be proved theoretically (Deb et al. 2002b). The reduction results by different approaches are shown in the following subsections.

5.2.2 Correlation analysis

The modified NCIE matrix plays an important role in the correlation analysis of our approach. The correlation matrix is another popular metric of correlation of multiple variables. We embed these two different matrices separately into our objective reduction approach to show the behavior of the modified NCIE matrix. The two objective reduction approaches are both embedded in NSGA-II. The differences

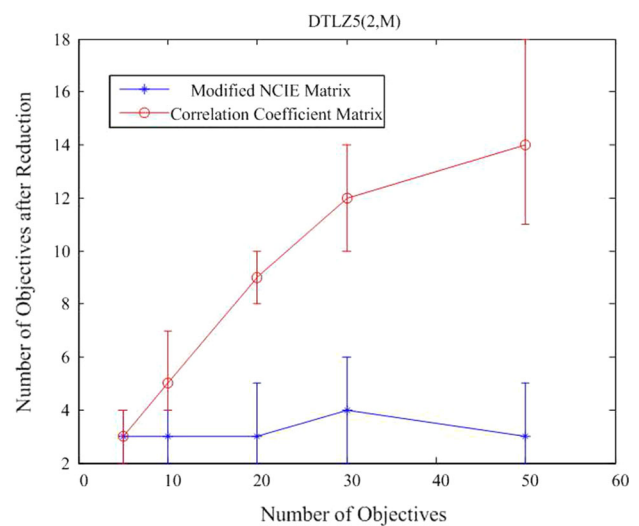


Fig. 4 Median number of objectives after reduction over 200 generations of our objective reduction approaches based on modified NCIE and correlation matrices on DTLZ5(2, M)

in reduction performance and execution time are summarized below.

The number of objectives after reduction over 200 generations is shown in Fig. 4. For the five-objective DTLZ5, both approaches based on the modified NCIE and correlation matrices reduce the problem to a three-objective optimization problem. However, when the number of objectives increases, the approach based on the modified NCIE matrix reduces more redundant objectives than that based on the correlation matrix. For the 50-objective problem, the approach based on the modified NCIE matrix reduces it into a three-objective problem, while the correlation matrix-based approach can only reduce the number of objectives to 14.

To investigate the differences of the modified NCIE and correlation matrices further, we show the number of times of each objective retained over 200 generations after objective reduction in Fig. 5. The approach based on the correlation matrix cannot reduce DTLZ5 to a two-objective problem except for five-objective DTLZ5. For the five-objective DTLZ5, the approach based on the modified NCIE matrix retains $\{f_1, f_5\}$, which is not the best reduction result. However, when the number of objectives increases to 50, the chance of retaining $\{f_1, f_M\}$ by the approach based on the modified NCIE matrix decreases, but that of retaining $\{f_{M-1}, f_M\}$ increases. According to Fig. 5, the modified NCIE matrix performs better than the correlation matrix on keeping objectives when the number of objectives is large, e.g., 50.

Figure 6 shows the execution time of NSGA-II with our objective reduction approaches based on modified NCIE and correlation matrices on DTLZ5(2, M). With the increasing number of objectives, both approaches increase their exe-

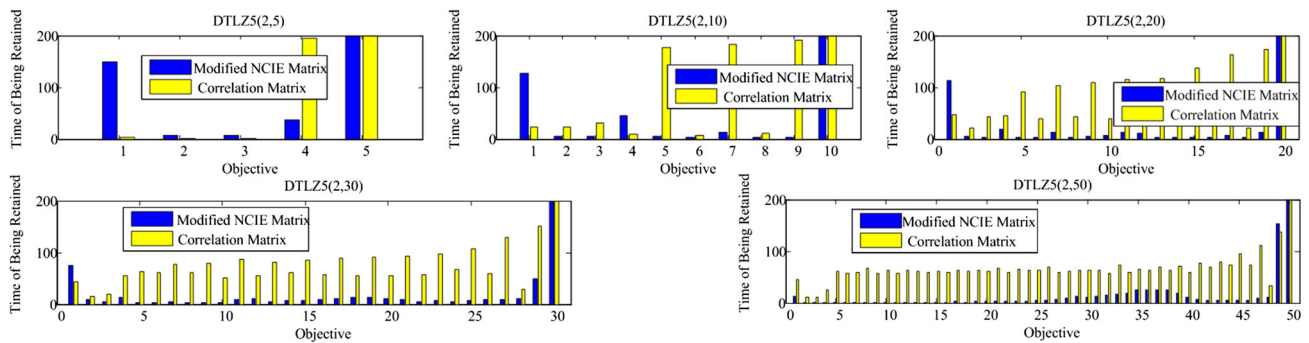


Fig. 5 Number of times of each objective retained over 200 generations after our objective reduction approaches based on modified NCIE and correlation matrices on DTLZ5(2, M)

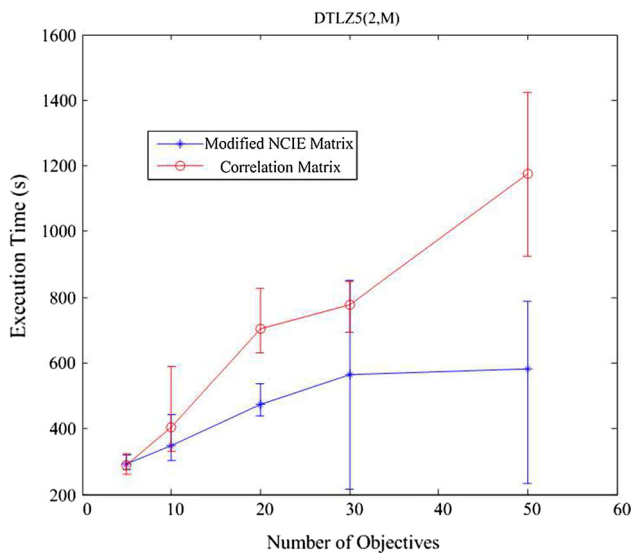


Fig. 6 Execution time of NSGA-II with our objective reduction approaches based on modified NCIE and correlation matrices on DTLZ5 (2, M)

cution time. However, the approach based on the modified NCIE matrix increases its execution time much slower than the approach based on the correlation matrix.

From the above results, we find that the approach based on the modified NCIE matrix performs better than the approach based on the correlation matrix for the problems with a large number (e.g., 50) of objectives, because the correlation matrix cannot provide clear correlation information for the subsequent objective selection.

5.2.3 Classification for objectives

The classification of objectives plays an important role in our approach. We compare our approaches with different T s and the classification method. The experiment is conducted on the reducible and irreducible problems (DTLZ5(2,10) and DTLZ2 with 10 objectives). Our objective reduction

approach is embedded in NSGA-II. For the reducible problems, our objective reduction approach aims to reduce the most redundant objectives. For the irreducible problems, our objective reduction approach aims to keep the right correlation of objectives. Therefore, we adopt the median number of objectives after reduction over 200 generations to evaluate the behavior of our approach.

Figure 7 shows the number of objectives after reduction over 200 generations in 30 independent runs with different T s. As the classification method has no T , which cannot be shown on the horizontal axis as other fixed T s, hence we show it on the horizontal axis by a special position outside the interval $[0, 1]$. Comparing the two sub-figures, we find that the size of T affects the behavior of our approach on reducible problems more than irreducible problems. Our approach decreases its performance with increasing T . The classification method obtains the best performance of objective reduction. For the globally irreducible problem DTLZ2, there are still some locally redundant objectives and our approach reduces three objectives. If T is set too small, some conflicting objectives are considered as redundant objectives, which would lead to the wrong dominance structure. If T is set too large, some redundant objectives would not be removed, which would waste the computational expense. The suitable value of T varies across problems. Therefore, a robust classification is very important to our approach.

5.2.4 Objective selection

We adopt a direct method to select the most conflicting objectives according to the obtained NCIE matrix from correlation analysis. The objective reduction approach based on PCA (Saxena et al. 2013) is a well-known one. Therefore, we compare our objective selection method with the PCA method (using the same setting as in Saxena et al. 2013). As NCIE is a nonlinear metric, we also compare it with the kernel PCA (KPCA) method (with Gaussian kernel function). DTLZ5(2, M) ($M = 5, 10, 20, 30, 50$) is chosen as the test

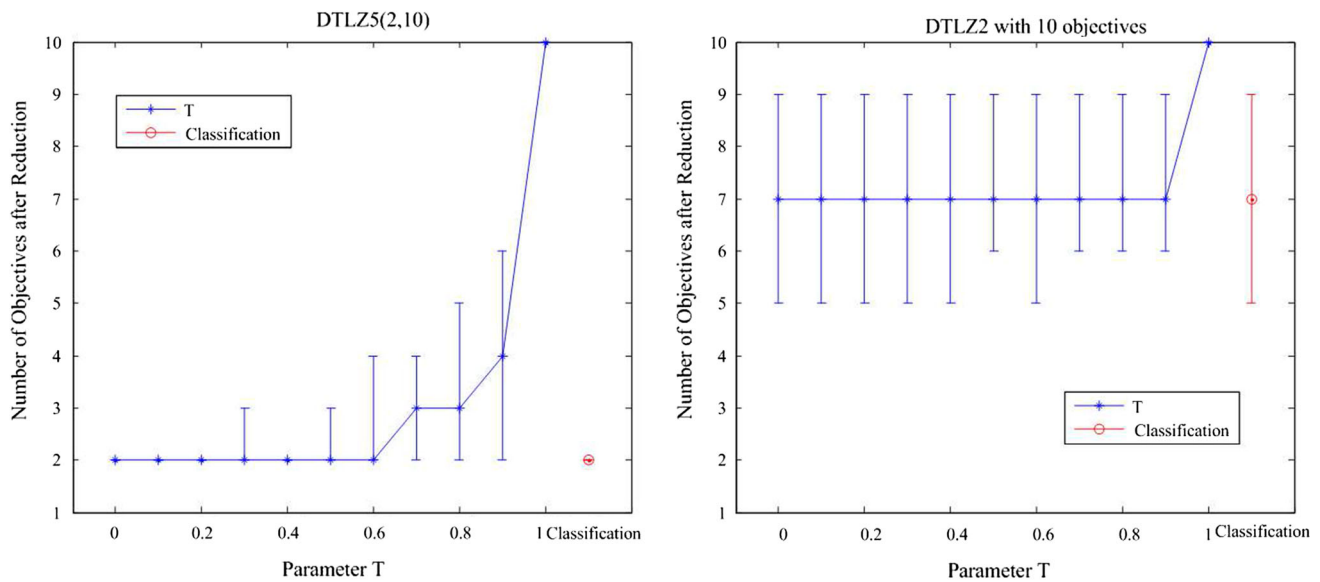


Fig. 7 Median number of objectives after reduction over 200 generations with different T s and the classification method

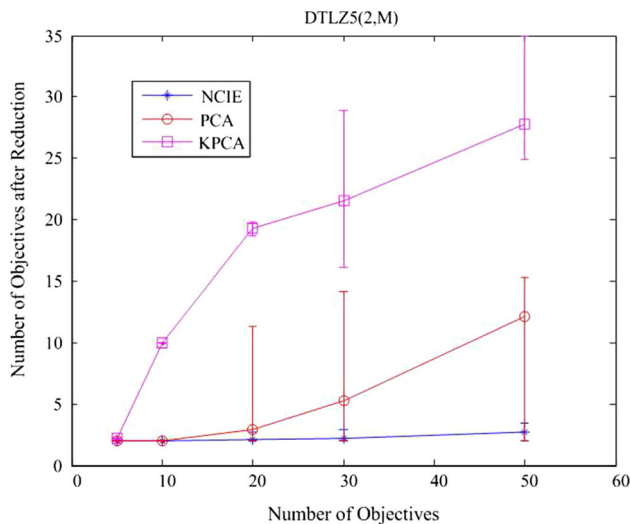


Fig. 8 Median number of objectives after reduction over 200 generations by our approach, PCA, and KPCA as the objective selection method on DTLZ5 (2, M)

problem in this subsection. We embed all the approaches in NSGA-II. The differences in reduction performance and execution time are summarized below.

The number of objectives after reduction over 200 generations is shown in Fig. 8. DTLZ5(2, M) can be reduced to a two-objective problem. All the approaches obtain two objectives for the DTLZ5 with five objectives. The KPCA-based method cannot reduce any objectives for the DTLZ5 with 10 and 20 objectives, but reduces the DTLZ5 with 30 and 50 objectives to 25 objectives. In contrast, our approach and the approach based on PCA have better performance. When the

number of objectives increases, our approach reduces more objectives than the approach based on PCA. For example, our approach reduces DTLZ5(2,50) to a three-objective problem, which is much better than the approach based on PCA.

In addition to the number of objectives after reduction, the obtained objectives affect the final results. We show the number of times of each objective retained after objective reduction in Fig. 9, from which we can find their different strategies of reducing objectives. As the approach based on KPCA does not have a good reduction behavior, the discussion is now focused on our approach and the approach based on PCA. For the DTLZ5s with 5, 10, 20, and 30 objectives, the approach based on PCA reduces the problems to $\{f_1, f_M\}$. Our approach reduces them to $\{f_{M-1}, f_M\}$ by more chances than other compared approaches. For 50-objective DTLZ5, the approach based on PCA obtains a three-objective set $\{f_1, f_{49}, f_{50}\}$, while our approach obtains $\{f_{49}, f_{50}\}$.

Our proposed approach can reduce objectives more efficiently than the approach based on PCA. This is reflected through two aspects, one is the number of objectives after reduction, and the other is the selected objectives.

Figure 10 is the execution time of these methods in 30 independent runs. For the DTLZ5 with five and ten objectives, our approach and the approach based on PCA use almost the same time. With the increasing number of objectives, the approach based on PCA requires longer execution time than our approach. Because of the poor performance of KPCA, few objectives can be reduced and its execution time is much longer than the other two methods.

We find that the objective selection in our approach performs better than the approach based on PCA in different aspects (objective reduction performance and execution

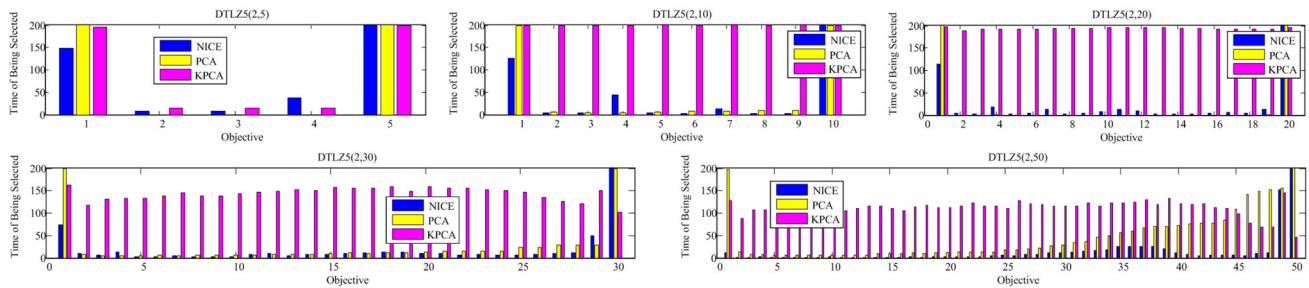


Fig. 9 Number of times of each objective retained after our approach, PCA, and KPCA as the objective selection method on DTLZ5(2, M)

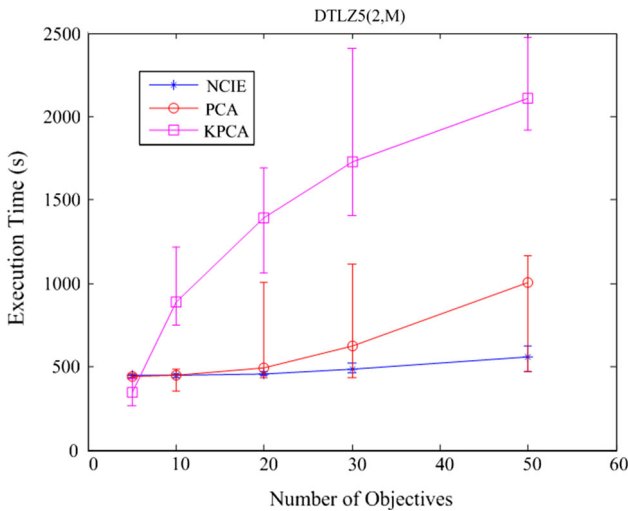


Fig. 10 Execution time of NSGA-II with our approach, PCA, and KPCA as the objective selection method in our objective reduction approach on DTLZ5 (2, M)

time). The main reason why our objective selection approach outperforms the approach based on PCA is that our approach reduces more objectives than the approach based on PCA. Because of the larger number of objectives obtained by the approach based on PCA, it cannot reduce the difficulty of the original problem.

KPCA, as a nonlinear method, maps the data to a high-dimensional space to keep its nonlinear characteristic. However, its kernel function has to be chosen in advance, which affects its performance significantly. From the results, we can find that the Gaussian kernel function is not suitable for DTLZ5.

5.2.5 Population size

Our approach uses NCIE to measure the correlation among objectives based on the population during the execution of MOEAs. To study the effect of the population size on the performance of NCIE, we embed our approach in NSGA-II with different population sizes (100 and 200). DTLZ5(2, M) ($M = 5, 10, 20, 30, 50$) is chosen as the test problem in

this subsection. We show the probability of each objective retained for 30 independent runs in Fig. 11. In the cases of 100 and 200 solutions, the retained objectives are very similar on all the tested DTLZ5 problems. In other words, the influence of the population size on the performance of NCIE is small.

5.3 Experiments on performance

In this subsection, we apply our approach to both Pareto-based and indicator-based MOEAs (NSGA-II [Deb et al. 2002a](#) and IBEA ($I_{\varepsilon+}$ -based) [Zitzler and Künzli 2004](#)). Both the reducible and irreducible problems are tested in the following experiments.

5.3.1 Pareto-based MOEAs

We embed our NCIE-based approach in NSGA-II on DTLZ1-5. The results are analyzed by Mann–Whitney U test ([Hollander and Wolfe 1999](#)). The significant ones are in boldface (the significant level is 0.05).

Table 3 shows the IGD values of the NSGA-II with our approach and the original NSGA-II on DTLZ5. DTLZ5 is a reducible problem. We find the original NSGA-II cannot solve the MaOPs with more than 10 objectives, whereas the NSGA-II with our approach can solve the MaOPs with 50 objectives. Our objective reduction approach reduces these problems into easier problems. As a result, the NSGA-II with our approach performs better than the original NSGA-II in most cases. However, the NSGA-II with our approach performs worse than the original NSGA-II on DTLZ5(7,10) because DTLZ5(7,10) has few objectives that can be reduced.

There are no objectives that can be ignored globally in the irreducible problems DTLZ1-4. To show the behavior of our approach clearly, the NSGA-II with random objectives reduced is also compared with the NSGA-II with our approach and the original NSGA-II on DTLZ1-4 in Table 4. After objective reduction, the NSGA-II with our approach can handle the problem with more objectives than the original NSGA-II. Comparing the performance of the NSGA-II with our approach and the original NSGA-II on five-objective DTLZ problems, the IGD values are not greatly improved by

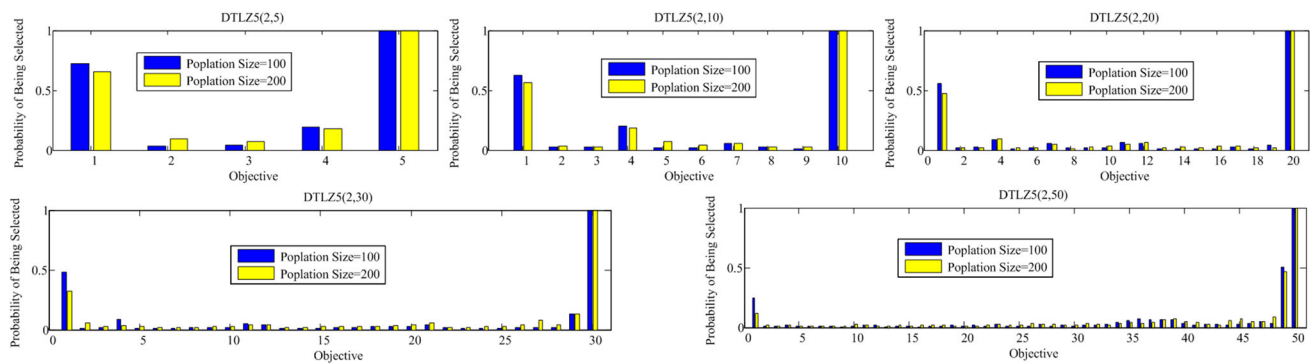


Fig. 11 Probability of each objective retained by our approach in different population sizes (100 and 200) on DTLZ5(2, M)

Table 3 IGD values of the NSGA-II with our approach and the original NSGA-II on DTLZ5 analyzed by Mann-Whitney U test

I	M	NSGA-II (NCIE)	NSGA-II	p value
2	5	0.0042 \pm 0.0000	0.0042 \pm 0.0000	0.0256
2	10	0.0042 \pm 0.0000	1.8023 \pm 1.5792	0.000
2	20	0.4613 \pm 0.2211	—	—
2	30	0.1086 \pm 0.2108	—	—
2	50	0.3886 \pm 0.3688	—	—
3	5	0.1155 \pm 0.1684	0.0550 \pm 0.0009	0.4735
3	10	0.0886 \pm 0.1068	3.4976 \pm 7.2347	0.0000
3	20	0.1192 \pm 0.1706	—	—
5	10	20.2226 \pm 38.9470	70.7366 \pm 47.8311	0.0003
5	20	134.9047 \pm 62.8815	—	—
7	10	93.9852 \pm 58.3535	67.9335 \pm 33.8466	0.0315
7	20	153.1658 \pm 54.6619	—	—

“—” Means that the algorithm cannot obtain the solution set within limited computational time. The significant results are in boldface (significant level = 0.05)

our approach except the hard-to-converge problem DTLZ1. Our objective reduction method did not seem to be effective on irreducible problems. Comparing the IGD values of the NSGA-IIs with our approach and random objectives reduced, the former outperforms the latter significantly. However, some objectives can be reduced locally; thus objective reduction to some extent promotes the population to better convergence. In short, the objective reduction for some irreducible problems can still make them easier for MOEAs by exploiting local correlation among objectives.

5.3.2 Indicator-based MOEAs

IBEA (Zitzler and Künzli 2004) is an indicator-based MOEA, which is well known for its ability for MaOPs. We embed our objective reduction approach in IBEA to analyze its effects on indicator-based MOEAs. The significant results are in boldface after being analyzed by Mann-Whitney U test (Hollander and Wolfe 1999) (the significant level is 0.05).

Table 4 IGD values of the NSGA-IIs with our approach and random objectives reduced and the original NSGA-II on DTLZ1-4 analyzed by Mann-Whitney U test

DTLZ	M	NSGA-II (NCIE)	NSGA-II (Random)	NSGA-II
1	5	3.5257 \pm 5.4437	29.7310 \pm 18.8620	28.1565 \pm 14.7707
1	15	15.8986 \pm 17.1624	36.6916 \pm 23.5109	—
1	25	17.7863 \pm 15.8425	27.6838 \pm 21.9433	—
2	5	0.5188 \pm 0.2108	1.2836 \pm 0.3187	0.4893 \pm 0.0741
2	15	2.1166 \pm 0.2912	2.2857 \pm 0.2768	—
2	25	2.4970 \pm 0.3190	2.7165 \pm 0.1811	—
3	5	129.0159 \pm 62.0312	209.6908 \pm 20.1481	160.7862 \pm 24.2639
3	15	216.4849 \pm 15.3258	232.9653 \pm 13.5670	—
3	25	230.2662 \pm 17.5408	277.0580 \pm 107.0687	—
4	5	0.7899 \pm 0.3023	1.1632 \pm 0.0114	0.5651 \pm 0.0613
4	15	1.2916 \pm 0.0634	1.3412 \pm 0.0399	—
4	25	1.3417 \pm 0.0602	1.4126 \pm 0.0750	—

“—” Means that the algorithm cannot obtain the solution set within limited computational time. The significant results are in boldface (significant level = 0.05)

Table 5 IGD values of the IBEA with our approach and the original IBEA on DTLZ5 analyzed by Mann–Whitney U test

I	M	IBEA (NCIE)	IBEA	p value
2	5	0.6529 \pm 0.1694	0.5784 \pm 0.1953	0.0764
2	10	0.6324 \pm 0.1789	0.5453 \pm 0.2052	0.0439
2	20	0.6734 \pm 0.1457	0.4877 \pm 0.1922	0.0003
2	30	0.6083 \pm 0.1835	0.5603 \pm 0.1601	0.2503
2	50	0.7018 \pm 0.1159	0.6298 \pm 0.1997	1.0000
3	5	0.7876 \pm 0.1536	0.7506 \pm 0.1633	0.9676
3	10	0.8257 \pm 0.1953	0.7815 \pm 0.1771	0.3369
3	20	0.7874 \pm 0.2094	0.8227 \pm 0.1836	0.9461
5	10	0.9366 \pm 0.1882	0.9312 \pm 0.2111	0.7764
5	20	0.8839 \pm 0.2212	0.8099 \pm 0.1878	0.4903
7	10	1.1193 \pm 0.1839	1.0622 \pm 0.2286	0.8817
7	20	1.1482 \pm 0.1584	1.0911 \pm 0.2212	0.7353

The significant results are in boldface (significant level = 0.05)

Table 5 is the result of the IBEA with our approach and the original IBEA on DTLZ5. For the reducible problem DTLZ5, no statistically significant difference is detected between the IBEA with our approach and the original IBEA, although our objective reduction approach seems to reduce the performance of IBEA.

As DTLZ1–4 are irreducible, the same random objective reduced IBEA as that in Sect. 5.3.1 is compared with the IBEA with our approach and the original IBEA. The results are shown in Table 6. There is no statistically significant difference among the three IBEAs.

The number of objectives appears to have little influence on the behavior of IBEA. In other words, the effect of our approach is not significant on IBEA for either reducible or irreducible problems.

5.3.3 WFG problems

WFG3 (Huband et al. 2006) can be reduced to a two-objective optimization problem, which is a different reducible problem from DTLZ5. We embed our approach in both NSGA-II and IBEA on WFG3 with 5–50 objectives. Table 7 shows the results of the MOEAs (NSGA-II and IBEA) with our approach and the original MOEAs in terms of IGD. The results are similar to those on DTLZ5 in Sects. 5.3.1 and 5.3.2. Our approach significantly improves Pareto-based MOEAs, but not indicator-based MOEAs.

5.3.4 Discussion

Generally, the advantage of IBEA is its good convergence ability on MaOPs, which NSGA-II cannot achieve. However, IBEA cannot obtain the results of good diversity because of the poor performance from $I_{\varepsilon+}$ (Hadka and Reed 2012). NSGA-II pays more attention to this aspect. Our experimental results support these two points. The aim of our approach is to improve the convergence ability while maintaining the good diversity of NSGA-II.

Our approach works well with Pareto-based MOEAs, but not with indicator-based MOEAs. Pareto-based MOEAs have difficulties on MaOPs because of ineffective Pareto dominance for large numbers of objectives. Our approach can reduce the redundant number of objectives to make the Pareto dominance work again. However, indicator-based MOEAs do not suffer from the Pareto dominance problem, even though the large numbers of objectives decrease their performance too.

For the reducible problems such as DTLZ5 and WFG3, our approach selects the most conflicting objectives, which decreases much computational cost. Thus, the NSGA-II with our approach can solve problems with more objectives, which

Table 6 IGD values of the IBEA with our approach and random objectives reduced and the original IBEA on DTLZ1–4 analyzed by Mann–Whitney U test

DTLZ	M	IBEA (NCIE)	IBEA (Random)	IBEA
1	5	1.8145 \pm 3.8061	3.7388 \pm 7.9257	3.7992 \pm 8.2498
1	15	3.9714 \pm 4.1883	3.5229 \pm 3.9579	7.8709 \pm 10.3937
1	25	13.0244 \pm 13.0977	8.4152 \pm 12.7078	12.6828 \pm 16.3969
2	5	1.1047 \pm 0.1011	1.1429 \pm 0.0667	1.0741 \pm 0.2317
2	15	1.3278 \pm 0.1319	1.2825 \pm 0.0649	1.2207 \pm 0.1260
2	25	1.3969 \pm 0.1818	1.3004 \pm 0.0977	1.3148 \pm 0.0318
3	5	43.9396 \pm 37.2782	39.3212 \pm 38.0931	34.4497 \pm 30.4998
3	15	39.3778 \pm 16.4310	42.8500 \pm 19.0732	47.5738 \pm 30.2017
3	25	39.1774 \pm 21.4049	42.2101 \pm 20.6839	52.6554 \pm 17.5275
4	5	1.1659 \pm 0.0002	1.1662 \pm 0.0007	1.1658 \pm 0.0002
4	15	1.3340 \pm 0.0000	1.3340 \pm 0.0000	1.3340 \pm 0.0000
4	25	1.3652 \pm 0.0042	1.3662 \pm 0.0001	1.3671 \pm 0.0005

The significant results are in boldface (significant level = 0.05)

Table 7 IGD values of the NSGA-II with our approach, the original NSGA-II, the IBEA with our approach, and the original IBEA on WFG3 analyzed by Mann–Whitney U test

M	NSGA-II (NCIE)	NSGA-II	p value	IBEA (NCIE)	IBEA	p value
5	0.0731 \pm 0.0609	0.4337 \pm 0.0714	0.0000	4.5874 \pm 0.1647	4.6532 \pm 0.0869	0.1404
10	1.3057 \pm 1.0653	2.3687 \pm 0.4006	0.0003	9.7115 \pm 2.2619	9.5860 \pm 2.5905	0.9031
20	3.7720 \pm 3.1806	9.1938 \pm 0.5134	0.0000	19.9937 \pm 5.4725	14.3213 \pm 9.5895	0.0859
30	5.8213 \pm 4.9378	–	–	30.7959 \pm 8.6535	24.3803 \pm 13.2391	0.2393
50	5.6174 \pm 6.8068	–	–	51.7169 \pm 14.1403	32.2109 \pm 22.5460	0.0275

“–” means that the algorithm cannot obtain the solution set within limited computational time. The significant results are in boldface (significant level = 0.05)

cannot be solved by the original NSGA-II. For example, DTLZ5(2,50) can be reduced to a two-objective problem, and the computational cost is decreased to 4 % of that for a 50-objective problem. However, with the increasing I in DTLZ5(I, M), the NSGA-II with our approach decreases its convergence ability. This is because the difficulties of those problems are still high after objective reduction. For the irreducible problems such as DTLZ1-4, the NSGA-II with our approach can solve those with 25 objectives, whereas the original NSGA-II can only solve the problems with five objectives. Furthermore, the NSGA-II with our approach can also obtain slightly better results than the original NSGA-II, because our approach was able to capture and exploit local objective interactions.

6 Conclusion

Since the correlation among redundant objectives might be either linear or nonlinear, the existing linear objective reduction approaches have limitations. We have proposed a novel objective reduction approach based on NCIE, which can handle both linear and nonlinear correlations.

In our approach, we employ NCIE, a nonlinear metric, to measure the correlation among objectives. In addition, we use a simple objective selection method without any pre-defined parameter, which results in the robustness of our approach. The experiments on DTLZ5 in Sect. 5.2.4 shows that our approach can select the most conflicting objectives for reduction. Our approach can be embedded in any MOEA to reduce the number of objectives, as demonstrated by the experiment on NSGA-II and IBEA. The experimental results show that our approach improves Pareto-based MOEAs (NSGA-II) on reducible problems (DTLZ5 and WFG3), but cannot improve the performance of indicator-based MOEAs (IBEA). At the same time, our approach also improves the performance of Pareto-based MOEAs on the irreducible problems (DTLZ1-4) slightly, because the difficulty of the original problems decreases locally, which promotes convergence.

However, there are some disadvantages of the NCIE approach that we have to overcome in our future work. (1) The reduction performance of our approach on the prob-

lems with more than 20 objectives is still not ideal. (2) The improvement of our approach on indicator-based MOEAs needs to be strengthened. (3) It will be interesting to evaluate our techniques on hypervolume-based IBEAs. (4) It will be useful to apply our approach to search knee areas of MaOPs (Bechikh et al. 2011).

Acknowledgments This work was supported by the National Basic Research Program (973 Program) of China (No. 2013CB329402), an EU FP7 IRSES Grant (No. 247619) on “Nature Inspired Computation and its Applications (NICaiA)”, an EPSRC grant (No. EP/J017515/1) on “DAASE: Dynamic Adaptive Automated Software Engineering”, the Program for Cheung Kong Scholars and Innovative Research Team in University (No. IRT1170), the National Natural Science Foundation of China (No. 61329302), National Science Foundation of China under Grant (No.91438103 and 91438201), and the Fund for Foreign Scholars in University Research and Teaching Programs (the 111 Project) (No. B07048). Xin Yao was supported by a Royal Society Wolfson Research Merit Award.

References

- Bader J, Zitzler E (2011) HypE: an algorithm for fast hypervolume-based many-objective optimization. *Evol Comput* 19(1):45–76
- Bechikh S, Said LB, Ghédira K (2011) Searching for knee regions of the pareto front using mobile reference points. *Soft Comput* 15(9):1807–1823
- Ben Said L, Bechikh S, Ghédira K (2010) The r -dominance: a new dominance relation for interactive evolutionary multicriteria decision making. *IEEE Trans Evolut Comput* 14(5):801–818
- Brockhoff D, Zitzler E (2006) Are all objectives necessary? On dimensionality reduction in evolutionary multiobjective optimization. *Parallel Probl Solv Nat PPSN IX* 4193:533–542
- Brockhoff D, Zitzler E (2009) Objective reduction in evolutionary multiobjective optimization: theory and applications. *Evolut Comput* 17(2):135–166
- Coello CC (2005) Recent trends in evolutionary multiobjective optimization. In: Abraham A, Jain L, Goldberg R (eds) *Evolutionary multiobjective optimization. Advanced information and knowledge processing*. Springer, London, pp 7–32
- Cvetkovic D, Parmee IC (2002) Preferences and their application in evolutionary multiobjective optimization. *IEEE Trans Evolut Comput* 6(1):42–57
- Dai C, Wang Y, Hu L (2015) An improved α -dominance strategy for many-objective optimization problems. *Soft Comput*. pp 1–7. doi:10.1007/s00500-014-1570-8

- Deb K (2001) Multi-objective optimization using evolutionary algorithms, vol 16. Wiley, New York
- Deb K, Jain H (2014) An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: Solving problems with box constraints. *IEEE Trans Evolut Comput* 18(4):577–601
- Deb K, Saxena D (2005) On finding pareto-optimal solutions through dimensionality reduction for certain large-dimensional multi-objective optimization problems. *Kangal report 11*
- Deb K, Tiwari S (2005) Omni-optimizer: a procedure for single and multi-objective optimization. In: *Evolutionary multi-criterion optimization*. Springer, Berlin, pp 47–61
- Deb K, Pratap A, Agarwal S, Meyarivan T (2002a) A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans Evolut Comput* 6(2):182–197
- Deb K, Thiele L, Laumanns M, Zitzler E (2002b) Scalable multi-objective optimization test problems. In: *Evolutionary computation, (2002) CEC 2002*. IEEE Press, IEEE Congress on, pp 825–830
- Farina M, Amato P (2002) On the optimal solution definition for many-criteria optimization problems. In: *Fuzzy Information Processing Society, 2002. Proceedings. NAFIPS. 2002 Annual Meeting of the North American*, IEEE Press, pp 233–238
- Fleming P, Purshouse R, Lygoe R (2005) Many-objective optimization: an engineering design perspective. In: *Evolutionary multi-criterion optimization*, Springer, Berlin, pp 14–32
- Fonseca C, Fleming P (1995) An overview of evolutionary algorithms in multiobjective optimization. *Evolut Comput* 3(1):1–16
- Freitas AR, Fleming PJ, Guimaraes FG (2013) A non-parametric harmony-based objective reduction method for many-objective optimization. In: *Systems, man, and cybernetics (SMC), IEEE International conference on*, IEEE Press, pp 651–656
- Gal T, Hanne T (1999) Consequences of dropping nonessential objectives for the application of MCDM methods. *Euro J Operat Res* 119(2):373–378
- Giagkiozis I, Fleming PJ (2014) Pareto front estimation for decision making. *Evolut Comput* 22(4):651–678
- Gong D, Wang G, Sun X, Han Y (2014) A set-based genetic algorithm for solving the many-objective optimization problem. *Soft Comput* pp 1–19. doi:[10.1007/s00500-014-1284-y](https://doi.org/10.1007/s00500-014-1284-y)
- Hadka D, Reed P (2012) Diagnostic assessment of search controls and failure modes in many-objective evolutionary optimization. *Evolut Comput* 20(3):423–452
- Hollander M, Wolfe D (1999) *Nonparametric statistical methods*. Wiley-Interscience, New York
- Huband S, Hingston P, Barone L, While L (2006) A review of multi-objective test problems and a scalable test problem toolkit. *IEEE Trans Evolut Comput* 10(5):477–506
- Hughes E (2005) Evolutionary many-objective optimisation: many once or one many? In: *Evolutionary Computation, 2005. CEC 2005. IEEE Congress on*, IEEE Press, vol 1, pp 222–227
- Hughes E (2007) Radar waveform optimisation as a many-objective application benchmark. In: *Evolutionary multi-criterion optimization*, Springer, Berlin, pp 700–714
- Ishibuchi H, Tsukamoto N, Nojima Y (2008) Evolutionary many-objective optimization: a short review. In: *Evolutionary Computation (2008) CEC 2008*. IEEE Press, IEEE Congress on, pp 2419–2426
- Ishibuchi H, Sakane Y, Tsukamoto N, Nojima Y (2009) Evolutionary many-objective optimization by NSGA-II and MOEA/D with large populations. In: *Systems, man, and cybernetics (SMC), IEEE international conference on*, IEEE Press, pp 1758–1763
- Ishibuchi H, Akedo N, Ohyanagi H, Nojima Y (2011a) Behavior of EMO algorithms on many-objective optimization problems with correlated objectives. In: *Evolutionary computation, (2011) CEC 2011*. IEEE Press, IEEE Congress on, pp 1465–1472
- Ishibuchi H, Hitotsuyanagi Y, Ohyanagi H, Nojima Y (2011b) Effects of the existence of highly correlated objectives on the behavior of MOEA/D. In: *Evolutionary multi-criterion optimization*, Springer, Berlin, pp 166–181
- Karahan I, Koksalan M (2010) A territory defining multiobjective evolutionary algorithms and preference incorporation. *IEEE Trans Evolut Comput* 14(4):636–664
- Khan S, Bandyopadhyay S, Ganguly AR, Saigal S, Erickson DJ III, Protopopescu V, Ostrouchov G (2007) Relative performance of mutual information estimation methods for quantifying the dependence among short and noisy data. *Phys Rev E* 76(2):026209
- Khare V, Yao X, Deb K (2003) Performance scaling of multi-objective evolutionary algorithms. In: *Fonseca C, Fleming P, Zitzler E, Thiele L, Deb K (eds) Evolutionary multi-criterion optimization, vol 2632., Lecture notes in computer science* Springer, Berlin / Heidelberg, pp 376–390
- Kim JH, Han JH, Kim YH, Choi SH, Kim ES (2012) Preference-based solution selection algorithm for evolutionary multiobjective optimization. *IEEE Trans Evolut Comput* 16(1):20–34
- Koksalan M, Karahan I (2010) An interactive territory defining evolutionary algorithm: iTDEA. *IEEE Trans Evolut Comput* 14(5):702–722
- Köppen M, Vicente-Garcia R, Nickolay B (2005) Fuzzy-pareto-dominance and its application in evolutionary multi-objective optimization. In: *Evolutionary multi-criterion optimization*, Springer, Berlin, pp 399–412
- Kukkonen S, Lampinen J, (2007) Ranking-dominance and many-objective optimization. In: *Evolutionary computation, (2007) CEC 2007*. IEEE Press, IEEE Congress on, pp 3983–3990
- López Jaimes A, Coello Coello C, Chakraborty D (2008) Objective reduction using a feature selection technique. In: *Proceedings of the 10th annual conference on Genetic and evolutionary computation*, ACM Press, pp 673–680
- López Jaimes A, Coello Coello CA, Aguirre H, Tanaka K (2014) Objective space partitioning using conflict information for solving many-objective problems. *Inf Sci* 268:305–327
- Ma X, Qi Y, Li L, Liu F, Jiao L, Wu J (2014) MOEA/D with uniform decomposition measurement for many-objective problems. *Soft Comput* 18(12):2541–2564
- Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P (1997) Multimodality image registration by maximization of mutual information. *IEEE Trans Med Imaging* 16(2):187–198
- McClymont K, Keedwell E (2012) Deductive sort and climbing sort: new methods for non-dominated sorting. *Evolut Comput* 20(1):1–26
- Pereda E, Quiroga RQ, Bhattacharya J (2005) Nonlinear multivariate analysis of neurophysiological signals. *Prog Neurobiol* 77(1):1–37
- Praditwong K, Yao X (2007) How well do multi-objective evolutionary algorithms scale to large problems. In: *Evolutionary computation, (2007) CEC 2007*. IEEE Press, IEEE Congress on, pp 3959–3966
- Purshouse R, Fleming P (2003) Evolutionary many-objective optimisation: an exploratory analysis. In: *Evolutionary computation, 2003. CEC 2003*. IEEE Congress on, IEEE Press, vol 3, pp 2066–2073
- Sato H, Aguirre HE, Tanaka K (2007) Controlling dominance area of solutions and its impact on the performance of MOEAs. In: *Evolutionary multi-criterion optimization*, Springer, Berlin, pp 5–20
- Saxena D, Deb K (2007) Non-linear dimensionality reduction procedures for certain large-dimensional multi-objective optimization problems: employing correntropy and a novel maximum variance unfolding. In: *Evolutionary multi-criterion optimization*, Springer, Berlin, pp 772–787
- Saxena DK, Duro JA, Tiwari A, Deb K, Zhang Q (2013) Objective reduction in many-objective optimization: linear and nonlinear algorithms. *IEEE Trans Evolut Comput* 17(1):77–99. doi:[10.1109/TEVC.2012.2185847](https://doi.org/10.1109/TEVC.2012.2185847)

- Sindhya K, Ruiz AB, Miettinen K (2011) A preference based interactive evolutionary algorithm for multi-objective optimization: PIE. In: *Evolutionary multi-criterion optimization*, Springer, Berlin, pp 212–225
- Singh H, Isaacs A, Ray T (2011) A pareto corner search evolutionary algorithm and dimensionality reduction in many-objective optimization problems. *IEEE Trans Evolut Comput* 15(4):539–556
- Van Veldhuizen DA, Lamont GB (1998) Multiobjective evolutionary algorithm research: a history and analysis. Tech. rep, Citeseer
- Wagner T, Beume N, Naujoks B (2007) Pareto-, aggregation-, and indicator-based methods in many-objective optimization. In: *Evolutionary multi-criterion optimization*, Springer, Berlin, pp 742–756
- Wang H, Yao X (2014) Corner sort for pareto-based many-objective optimization. *IEEE Trans Cybernet* 44(1):92–102
- Wang Q, Shen Y, Zhang J (2005) A nonlinear correlation measure for multivariable data set. *Phys D Nonlinear Phenom* 200(3):287–295
- Wang R, Purshouse R, Fleming P (2013a) Preference-inspired coevolutionary algorithms for many-objective optimization. *IEEE Trans Evolut Comput* 17(4):474–494
- Wang R, Purshouse RC, Fleming PJ (2013b) Preference-inspired coevolutionary algorithm using adaptively generated goal vectors. In: *Evolutionary computation, (2013) CEC 2013*. IEEE Press, IEEE Congress, pp 916–923
- Zhang Q, Li H (2007) MOEA/D: a multiobjective evolutionary algorithm based on decomposition. *IEEE Trans Evolut Comput* 11(6):712–731
- Zhang Q, Zhou A, Zhao S, Suganthan P, Liu W, Tiwari S (2008) Multiobjective optimization test instances for the CEC 2009 special session and competition. University of Essex, Colchester, UK and Nanyang Technological University, Singapore, special session on performance assessment of multi-objective optimization algorithms, technical report, pp 1–30
- Zitzler E, Künzli S (2004) Indicator-based selection in multiobjective search. In: *Parallel problem solving from Nature-PPSN VIII*, Springer, Berlin, pp 832–842